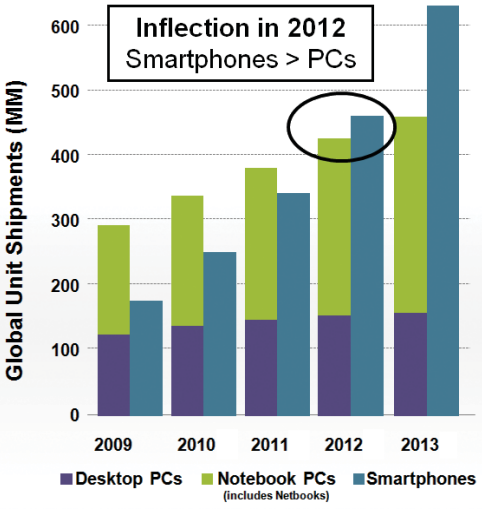
Сеть ЦОД для мобильной ИТ-инфраструктуры

В начале мая 2011 г. компания Extreme Networks анонсировала семейство коммутаторов для построения ЦОД нового поколения – Extreme Networks BlackDiamond X8 и Summit x670 с поддержкой стандартов VEPA, DCB, интерфейсов 40GbE, позволяющих создавать эффективные, высокомасштабируемые, мобильные и виртуализированные сети с низкими задержками и высокой энергоэффективностью ориентированными для облачных вычислений.

Сергей Гусаков – технический директор, компания Extreme Networks Россия и СНГ.

Введение

ИТ-мир в последние несколько лет стремительно меняется. Основными драйверами этих изменений является бурное развитие рынков облачных сервисов и мобильных устройств, прежде всего, смартфонов, которые уже стали мощным "мобильным" дополнением к настольным компьютерам и ноутбукам, и в 2012 г. по числу продаж будут лидирующей группой с объемом рынка более 50% (рис.1). При этом, по данным IDC, мировой объем рынка от продажи мобильных приложений в ближайшие 4 года будет расти с ежегодными темпами 60% и к 2014 г. он достигнет $35 млрд.



Ист.: *Katy Huberty, Ehud Gelblum, Morgan Stanley Research, сент. 2010.*

Рис. 1. Мировой объем рынка от продажи мобильных приложений в ближайшие 4 года будет расти с ежегодными темпами 60% и к 2014 г. он достигнет $35 млрд *(ист.: IDC)*.

Также стремительно растет рынок облачных сервисов. В соответствии с отчетами IDC (http://itmanagement.earthweb.com/netsys/article.php/3870016/IDC-Sees-Cloud-Market-Maturing-Quickly.htm), в 2009 г. на технологии, аппаратные средства и ПО, связанные с облачными вычислениями, было потрачено приблизительно $17 млрд, и к 2013 г., как ожидается, эта сумма увеличится до $45 млрд. Согласно исследованиям 451 Group and Tier 1 Research, объем мирового рынка облачных сервисов в 2013 г. достигнет более $16 млрд (рис. 2).

Результаты ответов на вопрос: "Сколько данных вашей организации в ближайшие 18 месяцев будет мигрировано в публичные, частные и гибридные облака?", по опросу, проведенному Network World в августе 2010 г., приведены на рис. 3.

Основным драйвером бурного развития отмеченных выше тенденций явилась виртуализация ИТ-ресурсов, в которой серверная виртуализация занимает ключевую роль. Именно она в значительной степени стимулировала появление новых стандартов и функционала на сетевом и уровне хранения данных в последнее время. Жесткая интеграция технологий виртуализации на всех уровнях сделала возможной появление облачных сервисов разных типов – SaaS (Software as a Service), IaaS (Infrastructure as a Service), PaaS (Platform as a Service), SaaS (Storage as a Service) и др., когда, прежде всего, корпоративные ИТ-услуги организациям стало возможно получать по требованию, не обладая при этом самой ИТ-платформой и системным/прикладным ПО, необходимым для их развертывания.

Следствием этих тенденций явилось: 1) полное размывание понятия физической границы датацентра с точки зрения получения ИТ-услуг; 2) необходимость поддержания очень высокого уровня автоматизации управления ИТ-процессами и качественно нового уровня информационной безо-

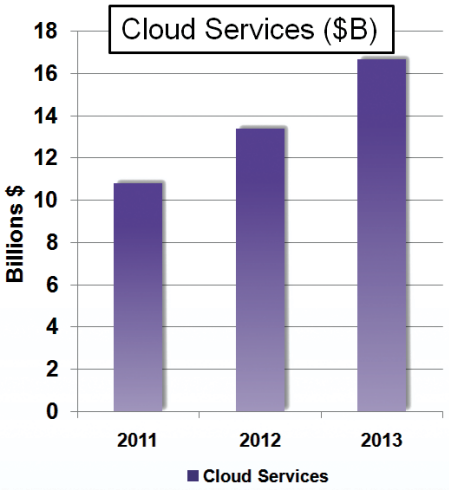


Рис. 2. Прогноз объема мирового рынка облачных сервисов (*ист.: 451 Group and Tier 1 Research*).

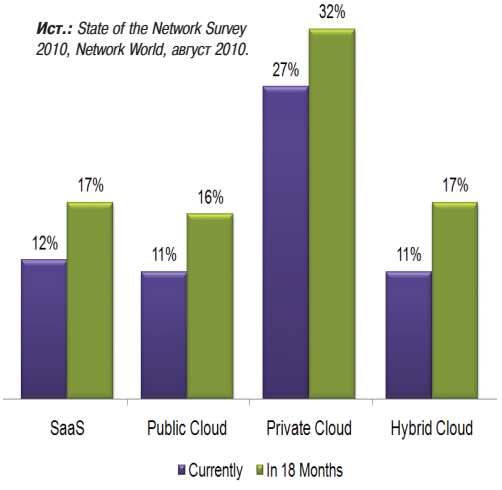


Рис. 3. Результаты ответов на вопрос "Приблизительно сколько данных вашей организации в ближайшие 18 месяцев будет мигрировано в публичные, частные и гибридные облака?", по итогам опроса проведенного Network World в августе 2010 г.

пасности в распределенной среде; 3) возможность развертывания и управления от тысяч до сотен тысяч виртуальных машин/серверов в рамках одного датацентра или нескольких консолидированных распределенных датацентров и необходимости существенного повышения эффективности системных администраторов.

В этих условиях критерии оценки эффективности и принципы построения датацентров существенно меняются, среди которых можно выделить следующие:

- *датацентры ранее состоящие из множества "островков" отдельных физических серверов преобразуются в консолидированный виртуализированный пул, безопасно разделяемый множеством арендодателей. Основным показателем становятся эффективность и мощность всего датацентра, а не отдельных его серверов/компонент;*
- *приложения, ИТ-сервисы должны развертываться "на лету", а масштабироваться прозрачно и просто без какого-либо физического добавления/замены устройств каждый раз;*
- *сетевая инфраструктура ЦОД уже не рассматриваемая как компонента для обеспечения связи между пользователями, серверами и хранением данных. Она должна быть глубоко интегрирована с уровнями серверов, СХД, средствами управления, чтобы обеспечить мобильность ВМ и облачных сервисов;*
- *за счет использования блэйд-серверов, многоядерных процессоров в современных датацентрах достигается очень высокая плотность компоновки (на одном сервере можно развернуть 64–128 VM вместо 8–16 ранее), что требует также и высокой плотности и производительности портов со стороны сетевой инфраструктуры;*
- *высокая плотность компоновки предъявляет гораздо бОльшие требования по энергоэффективности как с точки зрения экономичности, так и экологичности современных ЦОД.*

Extreme Networks Open Fabric – сетевая архитектура ЦОД нового поколения

В начале мая 2011 г. компания Extreme Networks анонсировала семейство решений Extreme Networks Open Fabric для развертывания датацентров, позволяющие организациям создавать более эффективные, высокомасштабируемые, мобильные и виртуализированные сети с низкими задержками и высокой энергоэффективностью ориентированными для облачных вычислений.

Архитектура Extreme Networks Open Fabric предназначена для построения сетей нового поколения, которые основаны на открытых стандартах для интеграции внутренних сетей ЦОД с сетями поставщиков облачных услуг, автоматически подстраиваются под динамично меняющийся ландшафт виртуальной среды, обладают пропускной способностью на порядок выше существующих сетей и позволяют использовать простую топологию из 1-2 иерархических уровней вместо 5-6-уровневой сети даже при наличии десятков тысяч подключаемых серверов. Аппаратной основой Open Fabric служит самый мощный коммутатор в истории Ethernet – Extreme Networks BlackDiamond X8, – обладающий пропускной способностью в 20 Тбит/с на коммутатор и 60 Тбит/с на стандартную стойку 19". При этом архитектура X8 не имеет теоретических пределов для дальнейшего масштабирования, а использование единой модульной операционной системы ExtremeXOS позволяет строить Open Fabric также и на уже существующих у заказчиков коммутаторах Extreme Networks.

Архитектура Open Fabric основана на открытых стандартах, что дополнительно подтверждается активной поддержкой OpenFlow и OpenStack. Учитывая постоянно усиливающуюся критичность сети ЦОД, только открытость сетевой архитектуры позволяет заказчикам выбирать и комбинировать наилучшие решения в единой сети, а также исключить чрезмерную зависимость функционирования компании от одного единого производителя решения.

Новый коммутатор BlackDiamond X8

BlackDiamond X8 – это ключевая компонента решений уровня enterprise нового поколения. Среди его основных характеристик:

- 768 портов 10GbE или 192 порта 40GbE неблокируемой коммутации в 1/3 телекоммуникационного шкафа (14,5 RU);
- общая пропускная способность коммутатора более 20Tbps;
- поддержка до 128 тыс. виртуальных машин в одном коммутаторе;
- один из самых низких показателей потребления мощности – 5 Вт на 10 GE порт;
- одна из самых низких задержек – 3 мкс для L2/L3 и порт-порт;
- охлаждение Front-to-back, позволяющее не нарушать схему горячих и холодных коридоров;
- одна из самых высоких пропускных способностей для трафика Layer 2/3/4 – до 11,4 млрд пакетов в секунду.

Новый коммутатор Summit X670

Summit x670 – это коммутатор Top-of-Rack со следующими характеристиками:

- 64 порта 10GbE или 48 портов 10GbE и 4 порта 40GbE неблокируемой коммутации в форм-факторе 1 RU;
- задержка менее одной микросекунды для коммутации Layer 2 и Layer 3;
- низкое энергопотребление и высокоэффективные блоки питания;
- модуль расширения на 4 порта 40GbE, программно настраиваемый как 16 портов 10GbE;
- охлаждение Front-to-back или Back-to-front, позволяющее не нарушать схему горячих и холодных коридоров вне зависимости от места размещения коммутатора.

Доступность коммутаторов серии BlackDiamond X8 для заказа ожидается в конце 2011 г. Модели коммутаторов Summit® X670, а также интерфейсные модули 40 Гбит/с для коммутаторов BlackDiamond 8800 доступны с июля 2011 г.

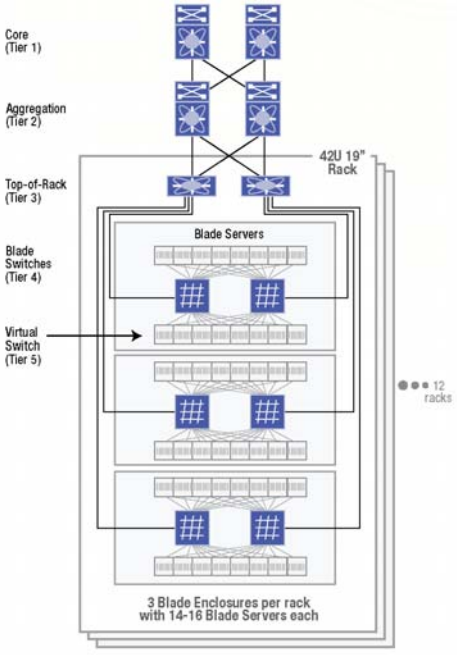
Технология Direct I/O и стандарт Virtual Ethernet Port Aggregator (VEPA)

Современная сетевая архитектура ЦОД для виртуализованных сред на базе стандартных серверов, в основном, представляется четырьмя или пятью уровнями (рис. 4):

- виртуальные коммутаторы (vSwitch);
- коммутаторы в блэйд-системах;
- коммутаторы, устанавливаемые на верху каждой стойки (Top-of-Rack, ToR);
- коммутаторы агрегации;
- коммутаторы ядра сети или ядра ЦОД.

В традиционной сетевой архитектуре без виртуализации и без блэйд-серверов сетевых уровней меньше на 2.

Типовой дизайн 5-уровневой сетевой архитектуры ЦОД для коммутации 576 блэйд-серверов



Дизайн Extreme Networks 2-уровневой сетевой архитектуры ЦОД для коммутации 576 блэйд-серверов

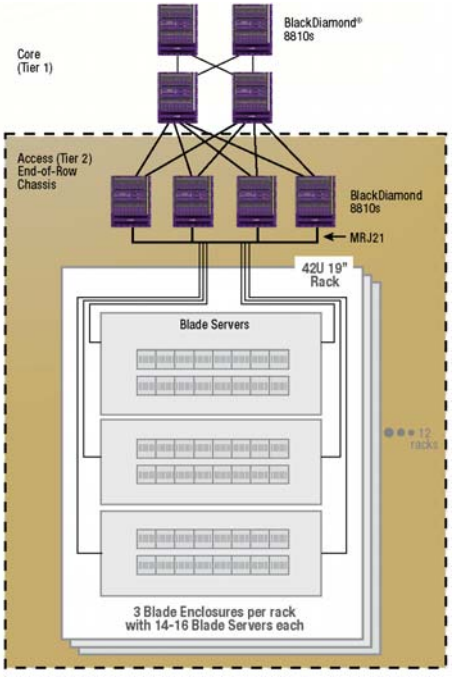


Рис. 4. Технология Extreme Networks® Direct Attach™ позволяет перейти от современной 5-уровневой сетевой (слева) архитектуры для виртуализованных сред к 2- или 3-уровневой (справа).

Дополнительные уровни – следствие развития ИТ-инфраструктуры. Уровень vSwitch (виртуальный коммутатор, 5-й уровень на рис. 4) появился в результате виртуализации серверов на базе, например, таких решений как: VMware, Microsoft Hyper-V, Citrix Xen Server и др. Это чисто программное решение, которое является уникальным для каждого гипервизора и разработано для того, чтобы обеспечить взаимодействие между: 1) виртуальными машинами, находящимися на одном физическом сервере; 2) ВМ и пользователями/приложениями.

Число vSwitch равно числу физических серверов, при этом каждый vSwitch должен конфигурироваться и управляться, что повышает сложность управления. Помимо этого, коммуникации на уровне vSwitch снижают информационную безопасность, т.к. не прозрачны для внешнего контроля и могут приводить к узким местам по производительности, т.к. трафик не контролируется.

Второй дополнительный сетевой уровень возник с появлением блэйд-архитектуры. С одной стороны она позволила существенно повысить плотность компоновки, упростить управление и развертывание физических серверов, а с другой потребовала введения отдельного коммутатора на уровне блэйд-шасси. Также к недостаткам блэйд-архитектуры, с точки зрения сетевой архитектуры, можно отнести:

– необходимость использования “переподписки” (использование одного канала несколькими источниками данных). Типичный блэйд-коммутатор содержит 24 порта. 16 портов используются для коммутации с блэйд-серверами, 8 – для обеспечения доступа на следующий сетевой уровень. Т.о. общее максимальное снижение производительности может достигать 50%;
– увеличение задержки, стоимости, затрат на управление, энергопотребление, требуемой мощности на охлаждение.

Развитие таких технологий, как Direct I/O и стандарта Virtual Ethernet Port Aggregator (VEPA), дает возможность коммутировать виртуальные машины на физическом коммутаторе, повышая производительность и упрощая менеджмент сети.

Подход компании Extreme Networks® к построению сетей ЦОД с помощью архитектуры Direct Attach™ – это следующая ступень развития виртуализации, позволяющая виртуальным машинам подключаться к сети напрямую, без необходимости коммутации на уровне серверного программного обеспечения. Такой подход к виртуальным ЦОД облегчает управление, мониторинг и поиск неисправностей, одновременно повышая сетевую производительность. Это достигается за счет того, что архитектура прямого подключения Direct Attach компании Extreme Networks позволяет производить обработку трафика виртуальных машин в традиционных и хорошо понимаемых сетевых инструментах, таких, как списки контроля доступа, межсетевые экраны и системы обнаружения/предотвращения вторжений.

Архитектура Direct Attach компании Extreme Networks взаимодействует с такими стандартами, как IEEE 802.1Qbg, что позволяет компаниям легко и прозрачно мигрировать на новый стандарт с сохранением сделанных инвестиций.

Технология Direct Attach компании Extreme Networks доступна путем добавления лицензируемого модуля операционной системы в коммутаторы ExtremeXOS®.

Решение Extreme Network Virtualization (XNV)

Традиционно сетевые администраторы имели ограниченный набор средств для поиска неисправностей и управления виртуальными машинами в сети. Обычно виртуальные машины создаются, настраиваются, перемещаются между серверными ресурсами и деактивируются динамически, с помощью инструментов администраторов серверного оборудования. Кроме того, для каждой платформы виртуализации требуется свой инструмент управления.

С появлением механизма XNV появилась возможность полностью унифицировать и передать "в пользование" сетевому администратору этот слой управления. Решение XNV позволяет полностью контролировать ВМ в течение их жизненного цикла на сетевом уровне.

Управление осуществляется через профили виртуальных портов – VirtualPortProfile (VPP), которые "привязываются" к отдельным виртуальным машинам и портам коммутатора. VPP дает возможность автоматически конфигурировать ACLs, QoS, а также позволяет автоматизировать задачи, касающиеся информационной безопасности, выделения необходимой пропускной полосы, сбора статистики всей сетевой инфраструктурой для виртуальных ЦОД. XNV обеспечивает автоматизированный трэкинг жизненного цикла виртуальных машин в сети, например, при миграции ВМ с сервера на сервер автоматически перемещается и VPP ВМ на соответствующий порт сетевого коммутатора (рис. 5). Также возможно назначение профиля VPP на любой порт коммутатора в режиме реального времени.

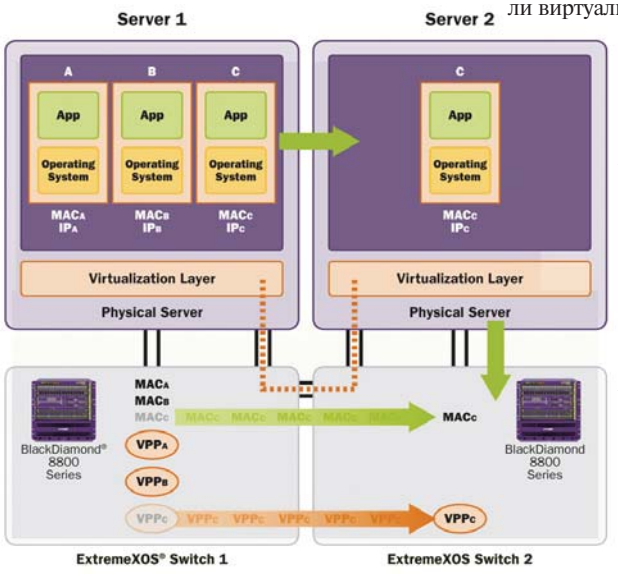


Рис. 5. Автоматическая перемещение профиля ВМ при ее миграции с одного сервера на другой.

XNV создает сетевую топологию с виртуальными машинами, с указанием их текущего месторасположения в сети, вплоть до порта коммутатора и с историей их перемещений. Это позволяет снизить ошибки конфигурирования, повысить безопасность и существенно снизить время поиска неисправности.

Настройка функционала XNV осуществляется через стандартный CLI (Command Line Interface) коммутатора. Кроме того, настройка также возможна через систему мониторинга и управления Extreme Networks Ridgeline™, существенно повышающую простоту и гибкость внедрения и использования функционала XNV.

Стандарт Data Center Bridging (DCB)

DCB – стандарт, который с помощью Ethernet нового поколения – Converged Enhanced Ethernet (CEE) позволяет осуществлять передачу сетевого трафика и трафика систем хранения данных (без потерь данных, подобно традиционно используемому для этого – FC) с помощью одного транспорта, что существенно позволяет снизить капитальные затраты на оборудование и операционные расходы на эксплуатацию. DCB поддерживает большинство протоколов для доступа к данным, включая NFS, CIFS, FCoE и iSCSI.

Разработаны следующие расширения для DCB:

– *EnhancedTransmissionSelection (ETS)* – основан на IEEE 802.1Qaz. Обеспечивает общую структуру управления для назначения/разделения полосы пропускания для разных классов трафика;
– *Priority-based Flow Control (PFC)* – основан на базе IEEE 802.1Qbb. Обеспечивает механизм управления приоритетами потоков на одном канале. Цель механизма – гарантировать минимальную задержку при передаче данных к/от СХД (позволяет потоку с более высоким приоритетом трафика продолжаться, в то время как более низкоприоритетный трафик может быть временно приостановлен);
– *Congestion Notification(CN)* – основан на базе IEEE 802.1Qau.Как и протокол PFC, обеспечивает управление перегрузками в сети, но более сложный, чем PFC;
– *DataCenterBridgingExchange (DCBX)* – протокол для обмена параметрами настройки, который используется для передачи настроек и конфигураций между узлами сети и обеспечения гарантии их непротиворечивости.

Multi-Switch Link Aggregation (M-LAG)

В современных Ethernet-сетях при увеличении числа виртуальных серверов, функционала, предоставляемого протоколом STP (Spanning Tree Protocol), становится недостаточно. Протокол STP воспринимает избыточные пути только в качестве резервных (standby) и не использует их для коммутации трафика. Кроме того, протокол STP накладывает существенные ограничения на масштабируемость сетевой инфраструктуры.

В коммутаторах Extreme Networks возможность полного использования всех путей сети достигается через функциональность M-LAG. Механизм M-LAG позволяет объединять два или более физиче-

ских канала от оконечных устройств или коммутаторов нижнего уровня в один логический канал на двух разных коммутаторах. При этом данный функционал работает на всех Layer 3 коммутаторах Extreme Networks, что позволяет использовать M-LAG в сетях любого масштаба.

Заменяя STP-архитектуру возможностями M-LAG, можно достичь удвоения утилизации сети на существующей инфраструктуре, избавится от сетевых проблем, связанных с протоколом STP, а также упростить конфигурацию, т.к. количество логических каналов существенно уменьшается.

Операционная система ExtremeXOS

Исторически сетевые коммутаторы и маршрутизаторы управлялись монолитной операционной системой (ОС). Однако современные конвергентные приложения, требующие передачи аудио/видео и цифровых данных, требуют от сетевой инфраструктуры гораздо более широких возможностей, которые являются вне досягаемости монолитных ОС. Конвергенция требует поддержания самых высоких уровней доступности и способности к поддержке быстро меняющихся бизнес-требований.

Монолитная ОС является статически откомпилированным программным образом, использующим отдельное адресное пространство. Программный модуль (или сетевая функциональность) не может быть динамически загружен или выгружен, потому что программные функции связаны в монолитной ОС. Динамически добавление, изменение или удаление какой-либо сетевой функциональности невозможно без перезагрузки ОС и времени простоя.

Немодульный дизайн требует временных издержек при добавлении новой функциональности (программных особенностей). Таким образом, монолитная ОС не позволяет достигать высоких уровней сетевой доступности, и ее поведение непредсказуемо в случае сбоев/отказов.

ExtremeXOS представляет из себя многоуровневую модульную ОС и, в сравнении с монолитной ОС, позволяет снять многие ограничения ей присущие. ExtremeXOS состоит из уровня аппаратной абстракции, ядра, модулей загружаемых в ядро, прикладных модулей и модуля управления конфигурацией (рис. 6). За счет такой архитектуры ОС удалось достичь следующего:

– *предсказуемости операций*. ExtremeXOS разработана, чтобы быть аппаратно независимой и обеспечивать непротиворечивые функциональные возможности на всех платформах коммутаторов Extreme Networks;
– *высокой доступности*. За счет устойчивости к программным и аппаратным сбоям, способности динамически останавливаться/перезапускаться, а также загружать/выгружать программные модули, не влияя на сетевые операции;
– *сетевой безопасности*. Защищенные на уровне оперативной памяти программные модули и безопасные транспортные протоколы защищают сетевую инфраструктуру через идентификацию, кодирование, целостность и DoS-защиту (Denial of Service);
– *расширяемых функциональных возможностей* За счет открытости (но при этом безопасной) к расширению сетевого функционала и возможности интеграции с пользовательскими или разработками третьих компаний на основе Extensible Markup Language (XML) и API;
– *простоты в управлении*. За счет: авторазвертывания пользователей или устройств и на основе Universal Port framework; гибкого управления политиками, используя Списки контроля доступа (ACLs); общего интерфейса ExtremeXOS, который может кастомизироваться через простой командный язык (Tool Command Language – Tcl).

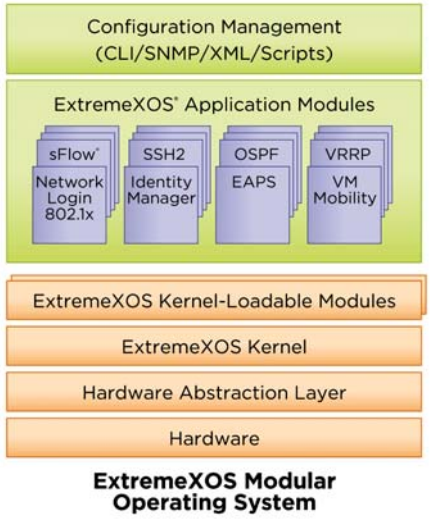


Рис. 6. Многоуровневая модульная архитектура ExtremeXOS.

Сравнение коммутатора BlackDiamond X8 с решениями, существующими на рынке

Коммутатор BlackDiamond X8 имеет очень большую плотность неблокируемых портов 10 и 40 GigabitEthernet. Так, всего в 14,5 rack unit можно установить до 768 портов 10 GigabitEthernet, или более 2300 портов 10 GigabitEthernet неблокируеммой коммутации в одном телекоммуникационном шкафу. На наш взгляд, в этом коммутаторе удалось получить максимальную плотность неблокируемых портов, существующую в настоящее время в индустрии (рис. 7). Кроме того, коммутатор BlackDiamond X8 не имеет теоретического предела по дальнейшему увеличению производительности т.к. не имеет шины Mid-Data Plane, т.е. линейные модули напрямую подключаются в модули коммутации, без использования шины, которая, как правило, является узким местом в коммутации. При необходимости подобрать решение на 768 портов неблокируемой коммутации 10 GigabitEthernet на оборудовании других производителей потребуется не менее 6-8 коммутаторов, подключенных между собой. Данная конструкция займет около 200 rack unit в телекоммуникационных шкафах вместо 14,5 RU, занимаемых одним коммутатором BlackDiamond X8.

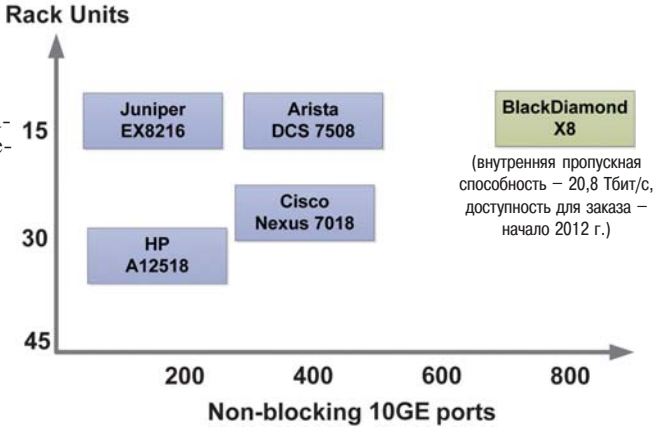


Рис. 7. Позиционирование BlackDiamond X8 с существующими на рынке коммутаторами *(подготовлено на основании официальной информации, представленной на сайтах компаний-разработчиков).*

Вместо заключения

Текущий 2011 год по праву можно назвать этапом революционных перемен в области сетевой архитектуры и сетей хранения данных, которые происходят не чаще одного раза в десятилетие. Концептуальные изменения выразились: 1) в переходе на конвергентные сети на базе lossless Ethernet на базе 10GbE и 40GbE (для коммутаторов BlackDiamond 8800 – доступность с июля 2011 г.), позволившие объединить на одном транспорте и платформе сетевые и SAN-инфраструктуры (законченные семейства решений уровня датацентра появятся к концу 2011 г.), а также значительно упростить и удешевить их развертывание; 2) в полнофункциональной аппаратной поддержке на сетевом уровне технологий серверной виртуализации, позволившей сделать возможной реализацию высокомасштабируемых гибких облачных сред. Эти нововведения заложили основу архитектуры сетевой инфраструктуры следующего поколения и десятилетия.

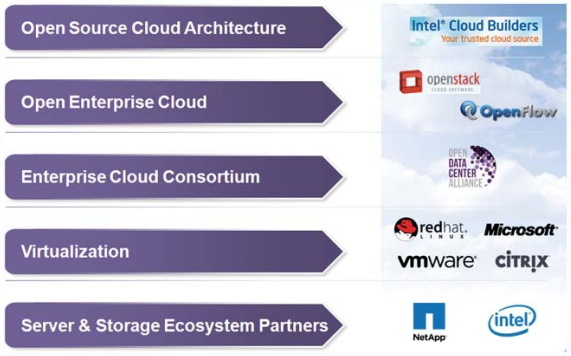


Рис. 8. Альянсы, с которыми работает Extreme Networks в области создания датацентров для облачных сервисов.

В начале мая 2011 г. Extreme Networks анонсировала членство в фонде Open Networking Foundation (ONF) для продвижения масштабируемых и открытых датацентров и облачных сетей. В настоящее время к инициативе присоединились 27 членов, среди которых основные поставщики сетевого оборудования, разработчики сетевого ПО и ПО виртуализации, поставщики чипов. Альянсы, с которыми работает Extreme Networks в области создания датацентров для облачных сервисов, представлены на рис. 8.

Сергей Гусаков,
технический директор,
компания Extreme Networks Россия и СНГ